1 Assignment 1 (100 points)

Object recognition and pose estimation is an active area of research in computer vision, so no one yet knows how to do object recognition in the general case. It is important to understand what we can do, and what we can't do, so that when your friend, your boss, or your customer comes to ask you for help, you know what you can and cannot give them. In this assignment, we will explore the power and limitations of using SIFT keypoints and descriptors for matching image features.

To submit your work, collect your inputs not provided as part of the assignment, the required outputs, and your answers to the written questions in pdf format (preferred) or Word, or .txt . Clearly label all of your work. For the math-oriented portion of the assignment, you may typeset in LaTeX or Word or provide a clear photograph of a hand-written solution. Work that is not legible will not be graded. Name a zip file `CS585_Assignment4_username.zip` and submit with web-submit. There is no code to submit for this assignment.

1.1 Learning Objectives

- Explore the types of image manipulations that SIFT can tolerate
- Explore the way that SIFT matching breaks down under image manipulation or change in 3D pose of objects
- Practice computing a histogram with weights and interpolation
- Review linear algebra that we will need when we discuss image mosaics the week following the midterm

1.2 Technical Task (50 points)

In Lab 4, you were provided with code that takes two images (an exemplar and a testing image), finds SIFT keypoints, extracts SIFT descriptors, and then performs matching using OpenCV's brute force matcher.

The brute force matching algorithm takes every descriptor found in the exemplar image and finds its nearest match in the testing image, *no matter how bad that match is.* You can use the slider on the top of the window so that the program will only display matches where the difference between the descriptors is less than an adjustable threshold. Collect input files and produce output files showing the feature matching working as described below.

(There are many different types of image descriptors (`http://docs.opencv.org/modules/features2d/doc/features2d.html`, `http://docs.opencv.org/modules/nonfree/doc/feature_detection.html`, and if you are feeling adventurous, you can try out some different ones. We will use FAST corners for creating image mosaics, since I have used them successfully in the past.)

1. (3 pts) Take a clear picture of some object in your environment and crop out a thumbnail of the object of interest. Use the program from Lab to perform matching of features from the thumbnail to the larger image. Save your output as "matchingYourObject01.png".
2. (24 pts)Using Photoshop, the GIMP, Instagram, iPhoto, OpenCV library calls, or any other image editing software, manipulate the larger image in the following ways. For each manipulation, create versions where you consider the feature matching successful, and where the feature

matching is unsuccessful. Save your outputs as "matchingYourObjectS02.png", "matchingYourObjectU02.png","matchingYourObjectS03.png", "matchingYourObjectU03.png",etc.

For our purposes, we will say that the feature matching is "successful" if you are able to find some matching threshold such that most of the features on the thumbnail are matched to something in the image, and most of those matches are on the corresponding part of the image (from whence you cropped the thumbnail), and not some other part of the image. The feature matching will be considered "unsuccessful" if you are not able to find such a threshold, or if you find that in order for many of the points from the thumbnail to match the corresponding part of the larger image, you must tolerate many spurious matches in other parts of the image. (3 pts each)

(a) Change the brightness of the image.

(b) Change the contrast of the image

(c) Change the size of the image (bigger or smaller)

(d) Blur the image

(e) Add noise to the image

(f) Draw some marks across the region of interest in the image to simulate occlusion

Based on your outputs, write a paragraph summarizing your findings. You should write approximately one sentence per image manipulation. (6 pts)

3. (23 pts)Take several pictures of the same object, crop a thumbnail from one of the images, and try to perform matching in one of the other images. You can use Part 3 from Lab 3 to capture a sequence of images from a webcam. Save your output as "matchingYourObject08.png","matchingYourObject09.png",etc.(3 pts each)

   (a) Use a planar (flat) object and rotate it in-plane (keeping the face of it towards the camera)

   (b) Use a planar object and rotate it out-of-plane (rotating the face of the object away from the camera)

   (c) Use a planar object and move it closer to and further away from the camera

   (d) Use a 3D object and rotate it in front of the camera

   (e) Use a 3D object and move it closer to and further away from the camera. See if you can catch images before the camera's auto-focus catches up with you.

   (f) Use a 3D object and have a friend turn the lights in the room on or off (the effect of this should be different than simply brightening a single image in photoshop, since the angle of the light will probably change.)

   Based on your outputs, write a paragraph summarizing your findings. You should write approximately one sentence per object manipulation. (5 pts)

1.3 Written Questions (20 points)

1. (3 pts)List four different types of invariance that we would like to have when matching feature descriptors (i.e. we would like to be able to match features even when images are different in what ways.)

2. (17 pts)Histograms are an extremely useful tool for summarizing information about a set of numbers. SIFT descriptors are just weighted histograms of the orientation of gradients in a window surrounding the keypoint. (Other uses are possible.)

   Below is a table of the orientations of some gradients in a 5 $\times$ 5 window. The magnitude of the gradient is also provided. Format:(orientation,magnitude)

(5°,1)	(−5°,2)	(−80°,1)	(−150°,4)	(−160°,5)
(−44°,1)	(−10°,2)	(−70°,1)	(−150°,3)	(−160°,5)
(−88°,2)	(−45°,2)	(−75°,1)	(−148°,3)	(−160°,4)
(−91°,3)	(−75°,1)	(−90°,2)	(−145°,2)	(−155°,4)
(−92°,3)	(−85°,1)	(−93°,2)	(−140°,2)	(−150°,3)

   (a) (5 pts)Using histogram bins of B=45 degrees [-180 to 180) construct a histogram of the orientation of the gradients by quantizing the orientation into bins using the floor operation (so an orientation of 89 degrees will go into the bin from [45 – 90), and an orientation of 90 degrees will go into the bin from [90 – 135)).

   (b) (6 pts)Using histogram bins of B=45 degrees, use linear interpolation to construct a histogram where the weight of the samples is divided according to the distance between the data point and the center of each bin.

   To do this, you should identify the two bin centers c_1 and c_2 on either side of the sample x, compute the distances between the sample point and the bin centers d_1 and d_2, and then divide the weight for the sample point using the ratio $(B - d_1)/B$ and $(B - d_2)/B$ (So, a sample whose value was 90 degrees would be divided equally between the bins for [45–90) and [90 –135), and a sample whose value was 22.5 degrees would be wholly assigned to the bin for [0 –45).)

   (c) (6 pts)Using histogram bins of 45 degrees, create a weighted histogram where each sample is weighted by the corresponding magnitude of the gradient using linear interpolation.

1.4 Lecture Preparation (30 points)

After the exam, we will start discussing image mosaics. To make image mosaics, we need to be able to represent the geometric relationships between source images comprising the mosaic. We will represent these geometric relationships using some linear algebra. To make sure that everyone is on the same page, here are some exercises. This linear algebra will not be on the midterm exam, but you will have an easier time in subsequent lectures if you are able to convince yourself of the truth of the following statements.

Assume that you have matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

and column vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

1. (2 pts)If values for A and y are given, state the protocol for finding the missing values of x in the equations $Ax = y$.

2. (2 pts)Prove that matrix multiplication does not commute for general 2×2 matrices of the form taken by A and B. (i.e. show that $AB \neq BA$.)

3. (2 pts)Show that the dot product $x \cdot y$ can also be written as $x^T y$.

4. (2 pts)Convert the following linear system into matrix notation (i.e. identify the corresponding coefficients for an equation of the form $Ax = y$) and solve the system (give the values for x).

$$\begin{aligned} 3x_1 \quad + \quad x_2 \quad &= 5 \\ x_1 \quad + \quad 5x_2 \quad &= 1 \end{aligned}$$

5. (2 pts)Is it possible to find a unique solution to the following system of equations? What is the constraint on the set of possible solutions?

$$\begin{aligned} 3x_1 \quad + \quad x_2 \quad &= 5 \\ 6x_1 \quad + \quad 2x_2 \quad &= 10 \end{aligned}$$

6. (2 pts)Using the matrices A and B, prove that $(AB)^T = B^T A^T$.

7. (2 pts)Using the matrix A and the vector x, prove that $(Ax)^T = x^T A^T$ (This is a special case of the previous question, but for a matrix and a vector (a matrix with only one column) instead of the more general case.)

8. (2 pts)Using the vectors x and y, prove that $x^T y = y^T x$.

9. (3 pts)Prove that $x^T A^T Ax$ is a scalar q

10. (3 pts)Remembering that the gradient is a vector containing the partial derivatives of a function (e.g. q) with respect to each dependent variable in the function (e.g. x_1, x_2), prove that the derivative of q with respect to x is $2A^T Ax$.

11. (3 pts)Prove that the derivative of $s = x^T y$ with respect to x is y, using the same idea of taking the derivative of s with respect to each element of x.

12. (2 pts)This somewhat unconventional rearrangement of variables will be very useful to us for finding the geometric relationship between two sets of points. Prove that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & 0 & 0 \\ 0 & 0 & x_1 & x_2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

13. (3 pts)*Homogenous coordinates* are obtained by tacking a 1 onto the end of a vector, like so:

$$x' = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}, y' = \begin{bmatrix} y_1 \\ y_2 \\ 1 \end{bmatrix}$$

Using this convention, many geometric relationships can be represented as matrix operations, including translation. (We would like to represent geometric relationships as matrix operations so that we can chain them together and manipulate them using the rules of linear algebra.) Find the result of $y' = Tx'$, where the matrix T representing a translation operation is

$$T = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$